

TITLE: NUCLEIC ACID INTERGRATION IN EUKARYOTES

Inventor: Hooykaas et al.

Docket No.: 2183-6028US

1/7

Fig. 1

Strain	LB' CAGGATATATTCAATTGTAAAT-CTC---CGA-GG	T-DNA	RB'	Chromosome, coordinate and location
WT.51	5' ⁻⁴ ATTGTATTATATTCAATTGTAAAT-CTC---CGA-GG 3'			XIV, 185311 (1 bp of target site DNA deleted), int. region
rad50k.1	5' ⁻⁶ TGTGGGTGTGATATTCAATTGTAAAT-CTC---CGA-GG 3'			XV, 1091277, tel. region
rad50k.5	5' ⁻⁷ GGGGGCATCAGTATTCAATTGTAAAT-CTC---CGA-GG 3'			XII, 465986, rDNA region
rad50k.6	5' ⁻²⁵ GAGGTAGATGTGAGAGAGTGTGTGGGTGTGAAGTCGA 3'			XV, 1091276, tel. region
mre11k.4	5' ⁻³ TCTGGTAGATATATTCAATTGTAAAT-CTC---CGA-GG 3'			XII, 459692/468829, rDNA region
mre11k.5	5' ⁻⁸ CACATATTCCTCATTCAATTGTAAAT-CTC---CGA-GG 3'			VII/X/XIII, 536090 OR 541678/472487 OR 483659/196667, LTR
mre11k.8	5' ⁻¹¹ CGACTACTTATATCCAATTGTAAAT-CTC---CGA-GG 3'			XIV, 6060, subtel. region
mre11k.11	5' ⁻⁷ GAAGAACCCATTATTCAATTGTAAAT-CTC---CGA-GG 3'			XIV, 4866, subtel. region
mre11k.14	5' ⁻⁷ TGGGTGTGGGTATTCAATTGTAAAT-CTC---CGA-GG 3'			VIII, 562588, tel. region
mre11k.17	5' ⁻⁹ TGGGTGTGGTGTCAATTGTAAAT-CTC---CGA-GG 3'			XII, 5727, subtel. region
trs2k.1	5' ⁻¹⁰ TGTGTGGGTGTGGTCAATTGTAAAT-CTC---CGA-GG 3'			IX/X, 69/52, tel. region
trs2k.17	5' ⁻¹ CGTCAAGGATATATTCAATTGTAAAT-CTC---CGA-GG 3'			XII, 1071797, subtel. region

Fig. 2

Sc	1	MRSVTNAFGNSGELM AKT GYRKED IL EGAFICELSETMFKESSYLEYKSPILLEI ES DEDEMSQL W RP
HS	1	MSGWESYYKTEGDEEAEPE Q EL AS GDY NS GW LS FL AS KAMFESOS I DEL T -PFDM S HO C OS Y ISK I ESD
At	1	
Sc	75	GT A CY F YC N DAKEGIYE E PT R ENATE M K Y ND L LED L SS G RISLYD F MF Q QT G SE K OR L SV I AT F EL D TEL
HS	80	RDL V V F Y G TE E K N SV N FK N W L Q D NP G AK R Y LELD D FK Q O G Q K R Y OD M GH G SD D Y S L Y EV L VC A N L F
At	1	
Sc	155	E Y PG C K Q Y SN K P F EL F T I D K Q A Q D -IDE R AR L R R Y LT I D L F D NK N F AT F F G Y A D K P F D I -EF Y S D I T O G SH
HS	155	S Y C F K S H A R E M L F T M E D N E B G N D -SA A S R A T R A G D I R T G I F P L P H L K K P G G -FD I H L F V K D I I S M E
At	16	RKG-SL K T E D K R N F L F T N E D E I G S M R I S V N E D M T R T L O X A K D A Q L I G I S T E I P L S O P D I O N I G F Y D I G L N S
Sc	230	T W E N T G L D E F D G P S E K P A K Y R S E N I L R N K E V K R I M E C P L I I D K T N F N G V K G T H E K A S V R Y K I E E H I R
HS	226	DE-DLRVHFEE S K L E E A R E T R K P A S R K L K L N K D -HS T S V G I T N L R O Z E K P P -PE K L Y R T N
At	93	DE-L E F M P S V G O K L E D -E D O E A V L A R K R I T F C E G -NS I E E G Y A L R P A P E S -E T L D S T N
Sc	310	QE A Y S K R E A N P I G -ED T G K I V H E P Y G D L D I N L E D S S D Q I V M E A Y T O K D A F L A E G I G S S K S H E Y E N N E D K S E F V
HS	294	EP V K A S T E T N T G G L H P S D I H S O I Y G S R G T I L E K E E T E I K R F D -DP G I M L G F K P L V -L R K H Y L P S L F Y
At	161	LP V K V E R S C I D T G -A H Q D P I O R H O P Y K N C I U F V E E L S O V K R I S -TR H L I L E G F K P L S -C L N D E H N L I P S H F Y
Sc	389	P A D A K E G S I R T L A L L A T R K R E A N I A L S G K I S N E H P S L Y T I S P S S -E K D Y N -EG F Y I V R N P F L D E I R E F S L
HS	370	P N E S L V I G S T L F S A L L I K C E R K V A A C R T P R N I P P Y E V A L V P Q D E E I S O K I Q V T P P G F A L V W L P F A D D K R F E
At	236	PS E K V I G S R A F I A L H S S M O L E R F A R A Y G -GT P P R L V A L Q E -E T S D G G Q V E P P G I N E E L P S A N D I R D E L
Sc	462	LSYDD G P E K Y D N M K V T O S I M G F N L D G C N P D E K N P I L O Z H Y S L H D Y L L -E T T F D E N E T P N T K K E P -
HS	447	-P E N E T E A T P O G G K M K A E D E L R F T -Y R S D S F E N P I L Q Q H E N L E A F A L D M E P O A V L T L P K V E R M N F P A E
At	313	HSK-PGVAXPRA D I O K K A S A M R S D L I D -ASVC O F A N P A L O H Y A L O A N D E E R T R E T L P D E E M N F P A E
Sc	537	H E D D S L E K Y I R N K I L E S K - K S E D P I I Q R L N K Y V K I W N - -MF Y K K F N - -DD N J S K E E K K
HS	520	S I L I D E F E L Y P P D N P E G K - V T K R I H D E G S K E P - -E V E Y S E E L K T H I S K G T L G K F T V P M L K E A R Y G L K S G K E
At	391	PA D O F K Q S E Y G D D P D E E I S SG A K E K S N K R K A D D G K Y Y I E L - AK T - G K L D L T V V E L T Y L T A N N L I E SG K E
Sc	593	PFD K W P K F N I - -
HS	598	E L E A T K H F D -
At	466	V I E N A L T H I G K -

Fig. 3A

Sc 1 MISALDSEPEPQNAPSPDFKWPCEELFVKIHEVQINGTAGGKSFSEKYVETISNIVEMWPKTVGNNIYPAVLVIALPYF
 Hs 1 ---
 At 1 -----VTEEIRES-----HLVSIFNWIQKSKTSSOKRSKFRFLDTYCKESDIFVAFR-----LIPESL-----F

Sc 81 MARRIYNKKEYVLIETTCSYKLPKNSATEORLKVW---QRVGKGGNLS--SLRVEETAKRRAEPPSKKTTIDVNVHYLDS
 Hs 11 ERMAYGIKEELAGIYVLLNLPRDSDKDALLLNTR--TPTGTHGAGDFIYIAYEVLRER-CYOKGSLTQEVNDLlds
 At 57 ERGSYGIKEEVVLAATCMLKAGCSDRDPALLNWRKGTTAKAGNAGNFSLIAEVLORRQGASGELTIKEENDLLDR

Sc 157 LEGDRFASGRGFKLVLKSKPFLHCEEMSEVIFEFFDILKLRVIGGOEKKULICDHPDAODYLSVLSDLKIVTSKLYC
 Hs 88 IASN---NSAKRKDIKEKE-LLCQLETOSSAIEOKWIRIILKDLKLGSQDIFESVEHNDAAFLHNVTEDIEKVCRCQLAC
 At 137 LASS---ENRKEENLVLISL-LEOKTN---AOEIKWIRIILKDLKLGSQDIFEFHPDAEDELENVTCDLKIVCEKLLC

Sc 237 PIVVRLKDDDSIKVGFCAFAPOLAKKINLSYEKICETLHDFWBEKIDGERIQHYMNYGHSIKIFSRGHDYTYLNGAS
 Hs 164 PSVGLSD--SISITFSASKEMLAA-INDHEIEKDQHOSBYIETKIDGERIOMHED-GEVVKYFSRNCVYIYDHOFGAS
 At 210 RIMORHKR--QDIEVGKAVEFOLAMATIDENAAWVPGKDVPAECKFDGERICHKN--GTDIHYFSRNEIDHNGYAHAM

Sc 317 LNSGK---OHLRFTDSNKECILDGEMETDAKRRVIEPGLVKGSAFEALSFNSINQVDFHPLYIIVFDMLYINGTSLTP
 Hs 239 PMEGS---EPFIHNAFKADQICILDGEMIAAPNTCTFLQKGLKTK---F---M---EDSDLQTCYCVEDVIMVNNKKLGH
 At 286 SDLIVQN-----ILVD---KCILDGEMEVDTSLRFAEFGSNOVIAKAAE-----EGLDSHKQLCYIAFDVLYVGDTSTIH

Sc 395 LEPLHOFVYILNSIPLKNIHEIVRS-----RCYEVESKKSLEVAISLGSEGAVINYYNSSSYNVSF
 Hs 313 EAIPLKYEHLSSIEPIPGRHEIVQK-----RQAHTKNIVDALNEAIEKREGEIGIKNQPLSIVKPDKF
 At 355 QSLKSERIEMLKKSKPLKGRDExVPEGGLNVHRPSGEPSWSIVVHAAAEVFFKETAVNRREGIVLKDELESKHEPEGDP

Sc 459 NNNWIKSKPEYIEEGFVNLDIENIGRDSGKDSFMLGLLVLDEEEYKKHOGDSSEIVDHISQEKHIONSERRVKKHISFC
 Hs 377 GEGWIKSKPEYISGLMFLDLDLIGGYGKGS-----RGGWISHFLCAZAKPPPGEFESV-EHFLS
 At 435 SGKWWIKSKPEYR-AGAIDLIIGGYGSGR-----RGGEIAOFLVAAAEAEANVYERR-EHISFC

Sc 539 SIEENGSSOEEFKLSDRATHGHWK---TSEVAPPASILEFG---SKIPAEWIL-PSESIKLEIKRSLSLDNTETNMQKVATNC
 Hs 438 RVGKCCMKELYSGGLKLAKYWKP-EHKKAPPSSILCGT-----EKPWVNLKPCNSMIVQIF---AAEIPSPMVKRTG
 At 495 RVGKCCSDRELNIVSKLIPYFNEHPKKAPPSSFYQVTHSKDRPMVWIDSPEKSIILSITS---DIRTERSASSVAPY

Sc 614 TLYGGYCKIIRYDKENTECYTLIDLYESFTVKSNEVYQAERSQLG-----LIRKKRKELISDSFHQYRKOLPISNIF
 Hs 508 TLRFPRIKIRDDKENHECMTLDDLEGLRGKASCKLASESILYIGGDEPQSKKRAAPGKVKVIGIEHLNAPNINNNK
 At 572 TLRFPRIKIRYDKPWHEDYCAFVELVNNSNGTQKQFSESTQDNPKVNKSSKRGEKKNWSLPSIFIQTDASIK

Sc 688 GLLFEVLSIIVTEDTGIRIISALEKTIIVEHGKETYNILKRHSIGDVNLISCKEETECKALIDRG--YDVEHFNWLL
 Hs 588 IISNIEEDFECVSGTDSQPKPYLENRTAEEGGYTQONPG-----PDTYCAIAGSENIEVKNIIILSNKEDVSPAWLL
 At 652 KTSIESNIEYFISNVPRSHQLETFHKMVENGGKFSMNENN-----SVTICLAAESREGYQAAKPO--EDVHFSWLL

Sc 766 CIAYKQALKEPNYCFNOKMRAVAEKRHDCLGDSIENDISETKLSSAKSQLSLPPMGELEIDSEWIKFPLFLFSNRC
 Hs 662 CFTKTSFPPWOPPMIHKCPSTEEIFAFSYDCIGDSYIITIDINQKLEVSGKIVSNEQTPEETASLADLEYRYSWDC
 At 725 CCSRNNPPLPKYFPHDARFTELQDSDDESDSYBWDIDLEGLKQVLSNAKRS--EDSKSIDYYKMLCPEKRWSC

Sc 845 ---IAYVPRRNISHEDDILIKIKLFGGTDQQLCNLIIPYDDEPLI---EDCNP2THEKQIKASDTI-----
 Hs 742 PLSPFRRHTVIIDEYAVINMSTKNEGTIAIKALEFRHGKVVSOLAEGVSHFEGDHSIADFKAFRFTE
 At 802 ---LSCCVYIYPYSQTLSTEEEALLGIMAKRLMLELJAGKVENNLAM-ASHVNLAMAEEPLDFTLVSSEMEKR

Sc 913 ---PKIAFVVVAPEWVHSIENCMYFEEDEPFWNY
 Hs 816 ---RKFVQIKESWVTSIDCCEQEFNOY
 At 878 LLLKKPLVVVSSWATISDCE-EKLCEDVTERPKYMEESDTEESDKSEHDTTEVASQGSAQTKEPASSKIAITSSRGR

TITLE: NUCLEIC ACID INTERGRATION IN EUKARYOTES

Inventor: Hooykaas et al.

Docket No.: 2183-6028US

4/7

Fig. 3B

Sc -----
Hs -----
At 957 SNTRAVKRGRSSTNSLQRVQRRRGKQPSKISGDETEESDASEEKVSTRLSDIAETDSFGEAQRNSSRGKCAKRGKSRVG

Sc -----
Hs -----
At 1037 QTQRVQRSRRGKKAIGGDESDENDELGDGNNNVSADAEEGNAAGRSVENEETREPDIAKYTESQQRDNTVAVEEALQDS

Sc -----
Hs -----
At 1117 RNAKTEMDMKEKLQIHEDPLQAMLMKMFPIPSQKTTETSNRTTGEYRKANVSGECESSEKRKLDATDNTSVNAGAESDV

Sc -----
Hs -----
At 1197 VPPLVKKKKVSYRDVAGELLKDW

Fig. 4

Sc	1	MDYPDP-D---TERRHTTDNHYGYNENDPITGDDSKTFHESAMIAKNNVDMYQSGDLEHVNKPSKELVQVLKWLRL
Hs	1	MSTADALDDDENIEELIVATDHLGMEKDAARGNDIETVPLDETERLACENEVOFLULLGGDLFHENKPSRKLHTCLELRL
At	1	MSREDFSD---TQRSLVATDCHLGYMEKDEIRHEDSFFAKETCSAEERQVDFULLGGDLFHENKPSPTTLVKAENLRL
Sc	77	LC[CMGDKPC[LE[LSDP[SQVFH2DEFTSVNYEDPNFNISIPVFG[SGNHDDAEGDSL[LC[PDILHATGL[NHFGI[TE[
Hs	81	Y[CMGDI[RPVQFE[LSDQS[VNFG[SKR[PPWNY[PG[CN[UNISIPVFSI[HGNHDDP[GADALCA[DLIS[CA[GVN[HEG[SI[
At	78	[HC[NDKPVQF[LSQ[VN[ON-AEG[QVNYEDPFN[CG[PVFSI[HGNHDDP[GV[NL[SA[DLIS[AC[NLV[NY[FG[Q[LG[
Sc	155	--SEKIKEVPILEOKGSTK[ALYGL[AVRDERL[RA[FKD-G[AVT[EV[TRP[---G[EW[FLN[EC[HON[FG[HTNT[AF[PE[
Hs	159	--V[KID[SP[ALI[OKGSTK[ALYGL[GS[IE[DERL[RM[EVN-[KVTM[RP[KE[DE[---NSWFN[LE[HON[FG[KG[STNF[IP[
At	157	SGV[GQ[IT[LP[IL[KKGSTK[ALYGL[GN[IR[DERL[RM[EV[Q[TP[PE[EV[O[TRP[CG[CDV[S[WTM[EV[SHONR[VK[SN[PK[NA[SF[
Sc	228	QFL[PD[FLD[VIWGHEHEC[PNL[VM[PIK[NFD[Q[PG[SS[VAT[SL[CG[EA[O[PK[VE[FL[LI[RY[GE[AP[MT[PI[PL[ET[PT[FR[MK[
Hs	232	QFL[DD[FD[VIWGHEHEC[IA[PT[NE[Q[OL[FI[Q[PG[SS[V[TS[LP[G[EV[KK[HV[GL[RI[CRK[MM[HK[PL[HT[VF[Q[PE[
At	237	HF[LP[FLD[VIWGHEHEC[LP[Q[EV[SG[MG[FI[Q[PG[SS[VAT[SL[CG[EE[PK[HV[LL[LI[CP[Q[PT[PI[LT[VR[PE[YT[
Sc	308	SISI[OD[VPH[---R[PHE[---KD[AT[SK[LI[EV[EE[MI[RD[AN[EE[EV[O[LP[DD[G[EG[DM[VA[E[PK[PL[IR[LR[VD[YS[PS[NT[Q[SP[DI[
Hs	311	IV[LA[HP[DI[FP[Q[PK[VT[Q[TS[CK[E[EN[EN[---E[EP[PL[EN[SH[---PE[PK[PL[VR[LR[VD[YS[G[---E[
At	316	I[VL[K[DE[SD[DI[DP[ND[---Q[ST[LE[HL[DK[V[VR[NI[TE[KA[---E[KA[EN[RS[---E[KL[PL[VR[EV[VD[YS[G[---E[
Sc	384	QV[EN[P[RR[FS[HE[EV[G[HE[Q[PK[RS[PT[RS[KK[G[NG[CS[SD[RD[V[KE[LF[SE[SG[GE[EV[EV[LV[N[---D[LL[N[Q[Q[
Hs	374	EP[FS[VL[R[FS[Q[KF[V[DR[VA[NE[RD[MI[HE[RE[Q[EV[Q[NG[EE[IN[FG[Q[EV[---P[EG[TT[LR[VE[DL[VK[Q[Y[FO[TA[EN[Q[Q[
At	373	MT[IN[P[Q[RF[G[KQ[KV[G[VA[NE[PD[DI[ME[SK[AS[K[SE[AN[MD[ER[---P[EE[LN[Q[Q[NI[EV[AL[V[---AES[NE[KE[
Sc	460	SLL[PE[V[GL[NE[AV[VK[FV[DK[DE[KT[KE[FE[SH[EV[SNE[V[G[IL[ST[EE[FL[TD[DA[EE[---K[AL[IK[Q[VR[AN[SV[RP[PT[---P[HE[ND[
Hs	447	SL[LI[ER[G[GE[AV[Q[EF[VD[MI[EN[DS[LP[EL[VK[VE[---K[TC[---K[EL[ER[RI[MA[ED[K[DE[EV[VR[RE[TF[Q[KN[---T[NE[ED[
At	438	E[LP[VN[DL[D[VA[EN[FV[DK[DI[LA[FY[SC[VO[NI[Q[---E[TR[G[SL[AK[SD[AK[K[FE[ED[DL[L[VK[G[C[LE[ER[FL[K[DR[ST[RP[PT[G[
Sc	536	[H[---FAF[NG[GL[DS[FR[SS[N[RF[VR[---S[PD[IT[Q[HD[NE[SR[TH[IS[Q[SS[SK[PT[PK[---R[VS[---I[AE[KK[IP[
Hs	517	[P[---[R[E[---TR[AL[RS[Q[SE[SA[AF[SD[---D[LM[MS[SL[LA[Q[AND[S[DD[SI[SA[AT[N[K[G[RG[---R[G[R[G[RG[Q[NS[A[GG[---
At	512	SK[G[FL[ST[AT[LT[SE[NL[TK[G[SS[G[IA[NA[SE[SD[DE[DT[Q[O[GL[AP[PT[G[R[Q[SS[AT[HE[G[PT[R[G[---GR[G[K[AS[AM[Q[OT[
Sc	602	-AF[SD[ST[VI[---DA[NE[EL[ED[NN[ND[Q[DF[DI[DE[ND[IM[---V[ST[EE[ED[---AS[GL[LN[G[R[TK[K[RP[PT[AS[---I[AC[UR[G[AG[
Hs	591	-QR[G[R[LP[---T[RO[Q[PS[RN[V[---TK[N[SE[VI[---DE[DV[ED[IF[PT[SK[TD[QR[SS[ST[SS[SK[IM[Q[SV[SK[G[VD[FE[ED[DD[DD[DP[
At	590	LD[SS[SL[LI[Q[OS[OR[S[A[S[A[AA[KS[AS[T[IG[ED[DV[---PS[ME[VI[PE[DN[KP[PD[SS[SE[DD[DE[ST[FK[G[R[RP[---A[RG[G[GP[
Sc	674	[S[R[PT[---R[TH[---I[---L[G[S[LL[AK[K[---K[---
Hs	669	FM[N[TS[---L[R[---I[LY[LL[AL[EN[---M[Q[PT[G[---K[M[CY[KL[---R[VS[---SL[RF[
At	666	[A[G[TS[SK[R[G[K[---I[SS[SS[LN[RL[---L[E[SS[KK[DD[DE[ED[DD[DE[PT[E[K[KL[LN[KS[O[P[RT[B[N[Y[G[AL[ER[

Fig. 5

Sc 1 -----MSAIIYKNSICGIRSFDSNDRE-TIEFGPLTLIVGMNGFKTTIECLKYATTGDLPPNS-KGGVFIHDPKVA
Hs 1 MLIFSVRDMFAKMSIUCGIRSEFGIEDAKOYTFFPLTIVGPGAGKTTIECLKYCTGDEPFG-KGNFVFIHDPKVA
At 1 -----MSTEDKMLIKGIRSDPERRN-EMTFFPLTLIVGAGAGKTTIECLKSVCTGILPPNARSGERFIHDPKVA

Sc 72 GEKDRAQKLAFTSANGLNATVTRNQQLMKTTIFKTLLEGQVAINNS-GDPSILSRSLEDAVPLYLGVPKAIL
Hs 80 QETDRAQINLOFKDVNGELFAVORSEVCTOKSKKTEFKLEGVETRT-KF-GEKVLSLCAEHDREISSLGVSKAIL
At 73 GETETKAQIKLRFRAAGKDKCIRSEFOLTONASMEKAGESVLTQINPHTEGEVCLSYRCAEDRESPALIGVSKAIL

Sc 151 EYVIFCQEDSLWPLSEPSNLKKFDEIFQAMETKALANLKSNDMSVILLLNQSEELAHDKDEKAALIHPGLC
Hs 158 NNVIFCQEDSNVPLSEGKALKQKFDEIFSATRYIKALETLEQDQPTQKZEPQZELKYLNOYKAGACEGRKQIPSKE
At 153 ENVIEVHQESNWPLOPSTLKKFDIIFSATRYTAKALEVFKLNDQFQINKEKLENQTLKDAKYKRESDACD

Sc 231 TKGHOYNEEVSEKOLMEITEZDKKFISNODFOKILSKVENIKNTLELIS-DOVERLSMSIDPLPEKPLANILANE
Hs 238 AGMTSSKELVKSYENEELDPKLNKEEENLSKEMKLDNEAKALDSRKQKEDKISELEEKSEKEFOGTDDEOLNLIVNH
At 233 ESTESSINVOLEBESSECKDAEYKEMMLKDERKLQDOSIKTAERSKFKCOSOYAAPEBENEDTIEKLKWSKE

Sc 310 SSVEMKNNQLRDPETASSLKDRLOSSLISLNSLHRQGELJACKETYEKN-FNLLSTIAKADEFKFOGLSNIENSIMA
Hs 318 OPHTRKERKLVPCHEKINKESRLLIISLSELEQGHLCLQHHRQHIFARDSLIGSIAATOLEDGFERGPFSER
At 313 EERLALLGTAKPKEREVDTETTISIISIARANYHETISNLCTEAHMLLKNERDSTIONHYNLGNVPSTFSTE

Sc 389 QVNHEMSQFKAELISDLTDIDOFAKDIOKETNLSDLIKSITVDSONLEY-NKKDLSKHDSD-BELAENLKSEKSL
Hs 398 CJKNFHKLSSERO-ECEAKTAMCLMNDFAKEETLKQKQDIEIIRDKRERGLR-IIELASEHSSKRONEILKNEKYLQOLEG
At 393 VVNLNTNRKSRIGEDEMELDKKKSEMETALSTAWDCYDANDRWEHESOKRAKDEIEGISKRIEKEHEDSFEEI

Sc 466 EODSENHELENLTYKEKIQSWSEMSISPKLNQKIEEHNESIILNOQENFDRMKTNOQALYAKLIGTKSINTRL
Hs 476 SNDRELELDPELIAKEREFLSKAHRNSNGETKMEISIOLNEKADLSTHKLDOEHTONHHTTETDQDSTHDKADED
At 473 SEVDKOTDEREKOVOVLEKTHONSERGFESKIEORGEHESYLSERHINTMRDREVNAAGDAEDE-LITRIDECKDRIR

Sc 546 FELOKITEKLDNSRIROFPLTCEBORAFLMDEOZLFINMOKNIAINNKZHELDERYTNAIYMLNTIEKLQDNOKS
Hs 556 HOFENIKSPHSD-----EHTSLIGEPNKKOLEDPLSKS-----KEINQTRDIAKLNHASSSEONKNHINNLERKEEQ
At 552 GVLGRLLPPEKD-----MKREKCALRSIREYDDLSLKS-----EAEKEVNMLOMKIOBNNS-LFKHNKDTESRHYI

Sc 626 KEKSHQILSENSPEDCTIDSYNDLLEETEISYKTALENLKMJOTTLFENRNLIAERDSCYCNSRKE--NESFKSKL
Hs 628 LSSYEDKLFVCGSQDIESLREKEEIKSSKORAMLAGATAASOFITOLTEENOS--CCPCCVCFOTEAELOEAT
At 622 ESKPOALKOSETIDAPKLLISAKEKRDKREYNMANGMROQLEPEEKHAROEH-----CCCCERSIT-ADEEASPIK

Sc 704 LOEAKTETDANFEKILKDTSONEKEYLHSIRLLEKHHIITNSN-EKIDNSQFCLIKAKEEPKTSISKLDSEVDSTKE
Hs 706 DLQSKLHLAPLKLXTESEKKKSKRRDEMLGLAPLPROHIDLKEKEIPELRNKLONVNREHORLAND-EOETALLTIM
At 697 KORFNASSTTEKVALAVESSNAESVFOOLDKLRAPEEYSKLTTEIPLAETLQEHTEEGOKSEALDELGHSAOIE

Sc 783 DEKELAEESEERPLIEKFTYHEKELKDEHNSSKTESEELSIYNTSEDGICTVDELEDDQOREMNDSEELRKTESDLOEKE
Hs 786 PEESAKVC-TD-STIMTRFOMELKDIEKIAQQAALKQG-IDLEFTSCOVEOEKOKEHKLDESSKQBLNRKLIQOQ
At 777 APKKSTIAFLQOB-LENADREFOEVSYQDQEDHEYKLDFRGLVTHEEESLSSLSSKKEKJGEEAKERREDQHME

Sc 863 EKRENSRMINLNEKELTSEHESSTOKONNSHSRSKRENNDHDSRVRKLEARNISLKNKSEAOVSILKKNER
Hs 864 EOIQHLKSTTMEKSEKLOASTNLROROO-----EEOTVELSTEISLYRKEAKOISPLETTLEKFOQKNEEINN
At 856 RDISCLOARWHATEKAKAHTDRDSTK-----AEDPERIAEKSOLDLIVYLTAEAGPLSKEREOILSYNLTERN

Sc 943 IQVRNKOKTADTRLRFQTYINAVVDFEAACDFELOTTIKELELNK-----AQMLELKEOLILKSNEEERKAD
Hs 941 TSNHIAQKENDIKEKVNHGMYKSTERHIDGCDPYMKQGETEAK-----VIAQLSECEKKEKKNEDERLMQDDEDT
At 933 QYEELAKKKRHYQDWTWALKASYKINCFTPYDOLKGERLSDQQLQSCRAKNEEAGEURNIDLERN

Sc 1018 SNEEETNLKONLELIEKSOQHESETSRALVONS-EZERDKYQESPLRTRFEELSSENAGKLGEAKOLGQIDSLT
Hs 1017 CKIOERWLODNLTLRKRNEEBKESEEEGKOLKRE-OMOLOOKSEHOKLEENITNEKRNHNLAEGROKGYEEEIIE
At 1013 CDHLRRNEDNLNYPTTAKEEBHTRETESLEQPLNIGAAAEAIKELRREDELLSELNRCRGT/SVYESSIS

Sc 1097 EPLR-TDYKDIKEKMYKEWSELOQRSFVTDIDHIVSKALDSAIMKZGKNDLNEIIDELWARTYSGTDIDTINIRSDE
Hs 1096 KELREPCEDAEERKYSMIVBTTTEVNLKLDLHYYTLDOAIMKFBMKMEEINKIIRLWSTYRGODIEYIEIRSDA
At 1093 VENSOAQYKDIKTHEDQHOLQTTTEANKLDDRYNALDKAEMFHMKMEEINKIIRELWOTYRGODIDYIEIISDS

Sc 1176 VS----SIVKGRSYNYRVVUMYKQDVEALMRGRCSAGQKVLASIIIRLALSETFGANCGSIALDEPTTNLDEENIESLAI
Hs 1176 ENVNSAEDKRRNMYRVVUMYKQDTEALMRGRCSAGQKVLASLIIRLALAETFCILNCGIIALDEPTTNLDRENIIESLAI
At 1173 NG----AGTRSYSYVIMOTGDTETLMRGRCSAGQKVLASLIIRLALAETFCILNCGIIALDEPTTNLDGPNSESLAGAI

Fig. 6

		1	50
At	(1)	MIGVDSKSSSTTFL	ETMVESEKTKHTC
Hs	(1)		MERKISR
Sc	(1)	MSQLTEFISCI	PVVNEEQNEEDERGLCKIQIEDGAMLETLDENSLSGERI
Consensus	(1)		IETL E SKIS
		51	100
At	(28)	LRLEISGADPIFVKGT	WHN--SRFDISVTDGSSS
Hs	(8)	IHLVSEPSITHFL	QVSWEKTLESGEVITLTOGHSA
Sc	(51)	EKMLVSEG	TGIFSKSSFGINDLIFTGENIDEESKKYVWYELLKMLTGHK
Consensus	(51)	IKLLIS A IFLK SW	S F ISLTDG SA
		101	150
At	(60)	WICNATEEEV	AERAAQWDQPVSEYL-KLAEOYIIGFO
Hs	(43)	WTGTVSESEI	SOEADDMMAMEKGKYV
Sc	(101)	VYIASLDEKVVF	TKWTCRMQDEVWKVVMEL
Consensus	(101)	WTC MSEDEVA	A DLD S YI KLAELTHPV L
		151	200
At	(95)	OPNSVYSFSDALEGSKR	ISWTFEKEGKLEWRWKCKPSDDSKKLTVGIL
Hs	(79)	GPADVYTFNF	SKESCYFFF
Sc	(149)	KKGETDLE	FEMADKLYKDICC
Consensus	(151)	PADVYSF A E K I	NDKD S KE F K DD I RELLD
		201	250
At	(145)	ELMEANIRLSEEVY	NKTR-SFEKMRSEAERCLAQGEKLCDEKTEFESATY
Hs	(129)	YCLDTIAENQAKNEHLO	K-ENERLIRDWDVQGRFEKC
Sc	(199)	KLLETRDERT	TRAMMVTLLNEKKKKTRELHEILRONNIKLS
Consensus	(201)	FILLET E S MM	K E EKLIRE DILAQ EK LSDKD LESALY
		251	300
At	(194)	---AKFLSVLNA	--KKAKLR-----ALRDKE
Hs	(178)	---KFILVLNE	--KKTKR-----SLHNKLLNA
Sc	(249)	NTEVQKPISELNSPGKRM	ERKTVVEPQNLQKKLKDT
Consensus	(251)	KFISVLNA KK KIR	AL KL A RE
		301	350
At	(223)	---EESTDKAESFESGRS	DDEKSEEASKKATSSKARGGKRAARS
Hs	(215)	TAICSEM	TADRD
Sc	(299)	VIKMEDDDFDDFQFFGLSKRPIIT	AKDKLSEKYDDITSFGDDTQSISFES
Consensus	(301)	EDS DK FD SK DD A DDLS	KASAA SKGD S S S
		351	400
At	(265)		
Hs	(265)	DIAPSRKRRQRMQRNLG	TEPKMAPOENQLOEKEKPDSL
Sc	(349)	DSSSDVQKHLV	LEDNGIQISAGRSDEDYGD
Consensus	(351)	D A K	G A D D DS SA S
		401	423
At	(265)		
Hs	(315)	ENMSLET	LRNSSPEDLFDEI
Sc	(399)	HSEQSGND	REPCLQTESETDIET
Consensus	(401)	R	D